Optimization of Images Clustering Using Learning Automata

Mehdi, Sadeghzadeh
Department Computer, Mahshahr Branch, Islamic Azad University
Mahshahr, IRAN
e-mail: sadegh_1999@yahoo.com

Mona, Nasrollahi Nia
Department Computer, Dezful Branch, Islamic Azad University
Dezful, IRAN
e-mail:mona.nasrolahi@yahoo.com

ABSTRACT
Today, the management, supervision and categorization of images in various images, such as medicine, military, pharmacy, engineering images, and etc., have been considered by many scholars. On the other hand, a huge amount of information has led to manual reviews of the image contents and the presentation of a proper category to be time consume and complex, and face many errors. Data mining techniques such as clustering and learning methods have played a very important role in improving this process. In this paper, the optimization of clustering of images using the learning machine has been discussed. With simulations performed on the STL-10 standard images set, it was observed that the level of clustering accuracy to be 97% and the accuracy of the image categorization in the proposed method to be about 91.43% and this ratio respectively has improved by about 14.43%, 4.43%, and 6.43%, compared to other SVM, KNN and regression methods, which is remarkable.

Keywords
Image clustering, learning automata, k-means algorithm.

1. INTRODUCTION
Content-based image retrieval (CBIR) introduced for the first time by T. Kato in 1992 to describe the automatic retrieval of images from a database based on color and form characteristics. Content-based image retrieval systems use visual concepts to search images in large databases based on image interest. This technology incorporates various research contexts such as image segmentation, feature extraction, image representation, semantic mapping of characteristics, storage and indexing, measurement of distance or similarity or retrieval criteria, which makes CBIR systems a major challenge.

The content-based approach means that the query routine automatically uses image visualizations rather than application of the predefined information by image as labels, titles, or keywords. In CBIR, each image stored in the database has its own low-level characteristics, which are stored in the visual characteristics database and is compared to the query image characteristics.

The characteristics of color, texture and shape are three main features for content-based image retrieval that are extracted from images by various methods of features extracting and each image is retrieved by its own feature vector.

2. LITERATURE REVIEW
The input of this algorithm is n sample data and the k value that determines the number of output clusters. At first, k samples are randomly selected from all the samples. These samples will be known as k-cluster representations.

Each of the remaining samples will be a member of the cluster that one of these representatives (k samples) belongs to; in other words, with the help of criteria such as the Euclidean distance, we calculate the similarity of each of the remaining samples with k representations, and the intended sample become a member of the cluster which is closer to it. Afterwards, for each cluster, a new representation is obtained by computing the average members of the cluster. This process is repeated until a standard coverage for termination work. For example, this process can continue until center of clusters not changed [1].

A random Automata is defined as a quintuple SA≡ {α, β, F, G, Q}, where r is the number of Automata acts, u = {α1, α2, ..., αr} is the set of Automata actions, β = {β1, β2, ..., βr} is Automata input sets, F = Q × β → Q is a new mode production function, G≡Q → α is the output function that is mapping the current state to the next output and Q (n) ≡ { Q1, Q2, ..., Qk} is the set of internal states of the Automata at the moment n.

The α set contains the Automata outputs or actions, which Automata selects in each step one action from the r actions of this set to apply to the environment. The inputs set (β) specifies the Automata inputs. Functions F and G map the current input status to the next output (next action) of automata. If the F and G mappings are definite, Automata are called definite Automata. In the case where the F and G mappings are random, the Automata are called random Automata [2].

The learning Automata are divided into two groups of static and variable structure Automata. In random Automata with a static structure, the probability of Automata actions is constant. While in the random Automata with the variable structure of probabilities, the Automata actions occur in each repetition. In learning Automata that includes variable structure, the change of actions probability is performed based on the learning algorithm. Also, in learning Automata that includes variable structure, the internal state of the Automata Q is represented by the probabilities of the Automata actions. In fact, the Automata shows the internal state of the Automata Q (n) in moment n with the probability vector of the Automata actions P (n), which is given below.

$$P(n) = \{p_1(n), p_2(n), ..., p_r(n)\} \quad (1)$$

So that $$\sum_{i=1}^{r} p_i(n) = 1, \forall n, p_i(n) = \Pr ob[\alpha(n) = \alpha_i] \quad (2)$$

At the start of the Automata activity, the probabilities of actions are the same and are equal to 1/r (That r is the number of Automata actions.)

SVM is a type of algorithms that is used for classification and regression. The basis of the SVM classifier algorithm is the linear classification of data. In the linear division of data, we try to select a line that creates a more reliable distance between the samples and has more boundaries between the classes. To use this algorithm, the data must pre-labeled and then we send them as input to the algorithm. An algorithm based on inputs builds a classification model. The given model is obtained based on the input instruction of the algorithm and the algorithm is tested using another part of

the data. For example, seventy percent of the data is used to train the algorithm and thirty percent to test the algorithm. The advantage of these methods is high accuracy [3].

3. RESEARCH BACKGROUND

Relevance feedback was first used in the early 1990s in the content-based retrieval of multimedia, and in particular in CBIR. [4] [5] and [6] have reviewed the methods introduced in the context of content-based image retrieval in recent years. In the [7] research, a CBIR retrieval system has been developed, in which the color and the texture characteristics are extracted from images, and the system provided by the name WBCHIR that performs the retrieval based on the color histogram wavelet, the color and texture characteristics are extracted through the transformation of the wavelet and histogram and the combination of these characteristics is used to convert and represent the images on a new scale.
Research [8] presents a CBIR system based on relevance feedback in which the image and system by interact with each other improve high-level queries which are based on low-level characteristics. In the research conducted in [9], a new approach is presented in the relevance feedback, in which the similarity function is updated using the two sets of related and unrelated images. Some of the methods introduced in this context are used to select the best combination of feature extraction methods and the similarity function of evolutionary algorithms [10] and [11].Using classification and clustering based methods is one of the short-term learning implementation strategies. The system described in reference [12] has used the Bayes classification to separate the collection of related and unrelated images. In the [13] research, the k-means clustering algorithm is used to image clustering using the result of each step of relevance feedback in the image retrieval system. In the [14] research, a k-means clustering algorithm is used to increase the accuracy of retrieval in retrieval systems and query optimization in each retrieval cycle.

4. PROPOSED METHOD

As seen from the flow chart in Fig. 2, the dataset first comes in a grouped way into the proposed system to be clustered. After the data is inserted in the system, the features are separated from the images in order and entered into the next step for preprocessing. At this point, the data cleaning on the extracted features is performs and, finally, the data is prepared for future use and is verified as acceptable. After the data is prepared, all images are applied to the clustering algorithm and, with the help of the learning Automata, a primary clustering performs on the images. After applying the clustering, the relevant data are sampled and ultimately, about 80% of the data is separated as training data for generating the model by the SVM classification algorithm and 20% of the data is separated for testing and evaluating the proposed method. Therefore, clustering process is applied on all the captured images by integrating the gradient, and the cluster centers are optimized by learning Automata for the images can be better placed in their cluster. The SVM classification algorithm is then used to categorize images.
SIFT local characteristics [15] are stable in terms of rotation, the light intensity, and the scale variations, and in addition have high resolution. These characteristics can be used as the basis for the optimized local feature extraction method. This algorithm reveals many stable feature points. Due to the large volume of data in recognition systems, it is necessary to consider this feature. It should be noted that the characteristics of color, texture and edges are taken into account in this algorithm and in general form the characteristic of the shape.
The SIFT algorithm has so far used limited in feature extraction. Also, few of the work done in this regard have benefit from the good characteristics of the algorithm in an inappropriate way. The other drawback of the performed methodology is to ignore post-processing in this area. In this study, we will first introduce a new use of local characteristics of SIFT. For this purpose, the number of matching points (based on the relative nearest neighbor) between the two images is considered as a similarity criterion. In this research, SIFT algorithm is optimized by making changes to achieve maximum speed.
In order to extract the feature, an algorithm is chosen as the basis for which the feature vector is unalterable relative to rotation, light intensity, and the scale variations. All these features are available in the SIFT algorithm. This algorithm locates the angled points and obtains the gradient of each one. Then this algorithm converts these gradients into histograms and the feature vector is formed.
To locate stable points, at first the original image is smoothed (blurred) and the first level forms the space-scale pyramid. Then the image size is halved and again blurred 5 times, thus the pyramid completes. After this stage, the neighboring images are subtracted and each of images which has an image of it's up and down neighbors is selected. Each pixel of these images is compared with 26 pixels of the neighbor, and if between these values there is the minimum or maximum, it will be selected as the probable stable (candidate) point. It should be noted that the Difference of Gaussian is a very good estimate of the Laplace image, which is obtained at a very low computational cost and is shown in the figures below**.**

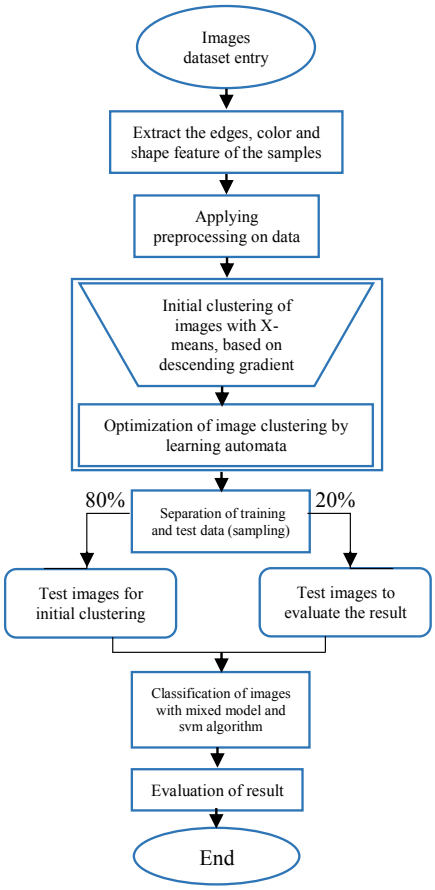

Figure 1. Suggested method

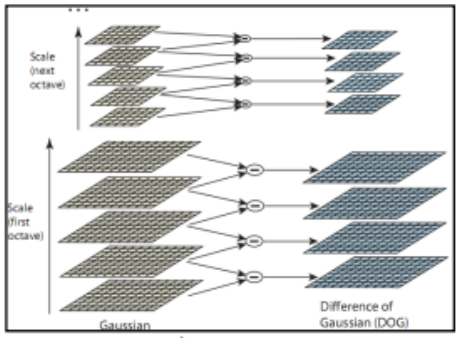

Figure 2. Space-scale pyramid. The Gaussian difference function is a good estimate of Laplacian

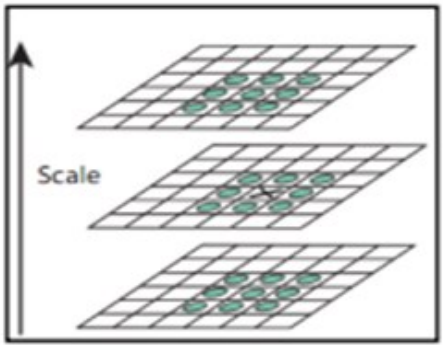

Figure 3. Selecting a stable key point based on my extremity over the key point neighbors

Direction assignment by calculating the gradient of each stable point, an angle is assigned to it as the direction of the stable point. All subsequent operators are converted depends on the calculated direction in this step and then applied to the image. Consequently, we achieve the irreversibility of the rotation. The calculation of the direction is obtained by the following formula.

$$\theta = \tan^{-1}\left(\frac{y_i - y_{i-1}}{x_i - x_{i-1}}\right) \qquad (3)$$

To construct a feature vector, 32*32 square pixels around each stable point are divided into 4*4 parts. In each section, a histogram of the angle and size values of each pixel is created. This histogram has eight sections that divide the angle between 0 and 360 degrees into 8 parts. Finally, with the addition of 16 × 8, the value of the range of the histogram parts of our 128-dimensional vector is obtained. This part is shown in the following figures.

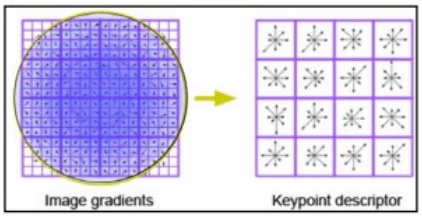

Figure 4. Putting together the values of the gradient domain around the key point, taking into account the Gaussian circular filter [15]

Figure 5. Feature extraction using the SIFT algorithm and the magnified image displayed for greater clarity [15]

In the figure above, it is seen that the corners of an image have been revealed as stable key points. In this figure, gradients around the image are displayed as flash, so after the features are extracted, preprocessing operations are performed.

As shown in the flowchart of the proposed method, by entering new data and pre-clustered data, the learning Automata trains model which is related to itself and by generating rewards and errors for each instance, categorizes the new data one by one. Finally, a category is selected that has more rewards and fewer penalties. In the event that the number of rewards and penalties are the same, in this case for the image entered, a cluster is chosen that has the least error or has the most samples in it.

After the images arrived at the K-Means clustering algorithm and a cluster was determined for that, the corresponding image is clustered with the K-Means cluster optimization by the learning Automata. Once the learning Automata has been implemented, if the inter-cluster distance is small or intra-cluster distance is large, then Automata will receive a penalty. In the following relation, we have shown how the Automata rewards and penalties are.

```
if((IG(IG_index)<threshhold1)||
(OG(OG_index)>threshhold2))
    reward = reward +1;
elseif((IG(IG_index)>threshhold1)||
(OG(OG_index)<threshhold2))
    penalty = penalty +1;
end
```

In the above pseudo code, IG is the inter-clusters distance and OG is the intra-cluster distance. Threshold indicates the threshold limit. Penalty indicates the amount of penalty and reward indicates the amount of reward.

After the images are separated into test and training samples, training data is developed to the SVM algorithm to generate the model and models are created based on extracted features. Then the test samples are given to the generated models and the new samples are sorted.

The authors would like to thank various sponsors for supporting their research.

5. SIMULATION

In this study, in order to evaluate the results of the proposed model and the comparison with other methods the dataset of images has been used under the title of Toronto Data Collection. This dataset can be downloaded from the official website https://www.cs.toronto.edu/~kriz/cifar.html. Evaluation criteria include:

$$\text{Re}call = \frac{TP}{TP + FN} \qquad (4)$$

In the formula, the correctness of the TP parameter indicates the number of images that are correctly detected; and the FP parameter also indicates the number of samples that are not properly divided.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (5)$$

In the calculation of accuracy, TN represents the number of samples that are falsely categorized.

After the simulation performed by using the MATLAB programming tool and with the number of clusters from 3 to 12 clusters, the following results were obtained.

The analysis of the table 1 shows that when the number of clusters is 10, the results considered optimal and by increasing the number of clusters or reducing the number, the amount of accuracy and recalling of the proposed method decreases.

Table1. The final results of the proposed method for clustering images according to number of clusters

Number of clusters	Accuracy	Recall
3	86.67	80
4	93.33	74
5	94.44	75
6	95.45	80
7	88	85
8	87.37	77
9	93	90
10	97.02	90
11	93.33	88
12	85	80
Average	91.43	81.9

In figure 6 results of proposed method compared with some other methods.

It should be noted that in this paper, the comparison is made for 100 images in each category, and only the accuracy of the clustering is evaluated and compared. As can be seen, the improvement of the proposed method is about 1.1 times compared with the SVM, KNN and regression methods. In other words, the improvement of image classification accuracy and correctness using the proposed method with the help of learning Automata is about 14.43%, 4.43% and 6.43%, respectively, compared with SVM, KNN and regression methods, which is abundant. It should be noted that the SVM method mentioned in the figure above was without descending gradient, learning Automata, and clustering.

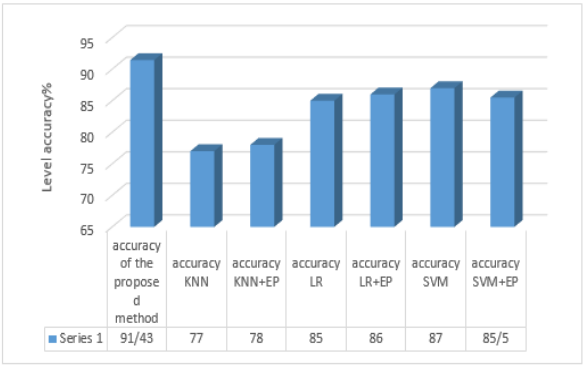

Figure 6. Comparison of the results of the classification of the images in the suggested method compared to other learning methods

6. CONCLUSION

Regarding the problems that the K-Means clustering algorithm has in terms of the number of pre-determined clusters, in this research, the X-Means clustering algorithm is used to optimize the number of clusters. By simulating the algorithm and the proposed model presented in this article, it was found that the number of clusters 10 has a better degree of accuracy and recall for categorization of images, and, accordingly, it is possible to separate the relevant data into test and training samples and eventually with the help of SVM algorithm attempt to categorize new samples. Finally, after simulating the proposed method and comparing it with other methods, it was observed that the accuracy of the proposed method is improved compared with similar methods.

REFERENCES

[1] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation", *IEEE transactions on pattern analysis and machine intelligence*, Vol. 24, No. 7, pp. 881-892, 2002.

[2] M. Esmaeilpour, V. Naderifar, Z. Shukur, "Cellular learning automata approach for data classification", *International journal of innovative computing, information and control*, Vol. 8, No. 12, pp. 8063-8076, 2012.

[3] S. Suthaharan "Support Vector Machine", *Machine Learning Models and Algorithms for Big Data Classification*, Springer US, pp. 207-235, 2016.

[4] R. Datta, D. J. Joshi, "Image Retrieval: Ideas, Influences and trends of the new Age", *ACM computing surveys*, 40, 2, 2008.

[5] P. B. Patil, M. B. Kokare, "Relevance feedback in content-based image retrieval: A review", *Journal of applications of computer science*, 10, pp. 41-47, 2011.

[6] Singhai, Nidhi, Shandilya, K. Shishir, "A survey on: content based image retrieval systems", *International Journal of Computer Applications*, 4, 2, 2010.

[7] M. Sinha and K. Hemachandran, "Content-based image retrieval using color and texture". *Signal Image Process. Int. J.*, 3, pp. 39-57, 2012.

[8] Y. Rui, T. Huang, T. Ortega, and S. Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8(5), pp. 644–655, 1998.

[9] A. Shamsi, H. Nezamabadi-pour, S. Saryazdi, "A new method in relevance feedback in content based image retrieval system", *Proceeding of CSICC*, (in Persian), 2010.

[10] C. D. Ferreira, J. A. Santos, R. S. Torres, M. A. Goncalves, R. C. Rezende, W. Fan, "Relevance feedback based on genetic programming for image retrieval", *Pattern Recognition letters*, 32, pp. 27-37, 2011.

[11] G. Raghuwanshi, N. Mishra, S. Sharma, "Content based image retrieval using implicit and explicit feedback with interactive genetic algorithm", *International journal of computer application*, 16, 43, 2012.

[12] M. Cord, P. H. Gosselin, "Image retrieval using long-term semantic learning", *IEEE International Conference on Image Processing*, pp. 2909-2912, 2006.

[13] J. Rejito, R. Wardoyo, S. Hartati, A. Harjoko, "Optimization CBIR using K-means Clustering for image retrieval", *International Journal of Computer Sciences and Information Technologies*, 3, 4, pp. 4789-4793, 2012.

[14] A. Aglawe, S. Bothara, Sh. Burhade, D. Chandak, "Clustering based CBIR system", *International Journal of research in computer science*, pp.2349-3828, 2015.

[15] D. G. Lowe, "Distinctive Image Features from Scale Invariant Key points" .*International Journal of Computer Vision*, 2004.